基于 SSA-XGBoost 模型的高精度密度预测方法研究

李瑞　吴文圣
（中国石油大学（北京）地球物理学院　北京　102249）

摘要　复杂岩性井段对密度测井数据精度要求很高，传统的计算模型不能很好的满足此时的高精度要求。为此提出利用机器学习回归预测模型提高密度测井曲线的精度，考虑到 XGBoost 模型的过拟合问题，基于 SSA 算法改进 XGBoost 进而提出了 SSA-XGBoost 密度预测模型。采用蒙特卡罗模拟双探测器密度测井仪器，获取不同密度地层数据以验证该模型的预测效果。利用 SSA 算法优化 SVR、RFR 和 LSTM 参数，构建 SSA-SVR、SSA-RFR 和 SSA-LSTM 模型预测模拟地层密度，并使用量化评价指标和泰勒图模型对比分析各个模型的预测性能。此外，还分析了不同预测模型对实际密度测井数据的预测效果。结果表明 SSA-XGBoost 模型的预测精度高于传统脊-肋图模型，在实际密度测井数据处理中具有广阔的应用前景。
关键词　SSA-XGBoost；蒙特卡罗模拟；机器学习；密度预测
中图分类号　TL99，TE19

Study on the Prediction of High-precision Density Based On the SSA-XGBoost Model

Li Rui　Wu Wensheng
（China University of Petroleum (Beijing), College of Geophysics, Beijing 102249）

Abstract　**[Background]**: Complex lithology well sections require high precision in density logging data, and traditional computational models cannot adequately meet the high precision requirements in these cases.　**[Purpose]**: This study aims to improve the precision of density logging curves using machine learning regression prediction models.　**[Method]**: Firstly, SSA algorithm was used to improve XGBoost, leading to the development of the SSA-XGBoost density prediction model. Then, Monte Carlo N-Particle transport code（MCNP） of dual-detector density logging tool instrument was used to obtain stratigraphic data of different densities to validate the predictive effectiveness of the model. By optimizing the parameters of SVR, RFR, and LSTM using the SSA, the SSA-SVR, SSA-RFR, and SSA-LSTM models were constructed to predict the simulated formation density. The predictive performance of each model was compared and analyzed using quantitative evaluation metrics and Taylor diagram models. Finally, the performance of different prediction models on actual density logging data was analyzed. **[Result]**: In the comparative analysis and processing of actual well density logging data with various models, the SSA-XGBoost model showed smaller errors between predicted and actual density, demonstrating high density accuracy and validating the precision of the method. **[Conclusion]**: The SSA-XGBoost model demonstrates higher predictive accuracy than traditional spine-ribs plot, showing great potential for applications in the processing of actual density logging data.
Key words　SSA-XGBoost, Monte Carlo simulation, Machine learning, Density prediction

---

第一作者：李瑞，男，1999 年出生，2022 年毕业于长江大学，现为中国石油大学（北京）硕士研究生，研究方向为地球物理测井
通讯作者：吴文圣，E-mail: wwsheng@cup.edu.cn

第一作者：李瑞，男，1999 年出生，2022 年毕业于长江大学，现为中国石油大学（北京）硕士研究生，研究方向为地球物理测井

通讯作者：吴文圣，E-mail: wwsheng@cup.edu.cn

随着国外内主力油田进入高含水开发阶段和致密油藏的发现，双探测器密度测井仪器精度较低，不能很好的满足复杂油气田开发需要[1-4]。提高密度测井的精度对识别油气层，计算地层孔隙度具有十分重要的重作用。传统的提高密度测井精度的响应模型主要有脊-肋图法（Spine-ribs Plot）、多参数联合反演（Multi-parameter Joint Inversion）和四维方程（Four-dimensional Equation）等，对测井数据的依赖程度较高，识别精度与效率较低[5-7]。随着硬件算力的显著增强，机器学习和深度学习算法持续取得突破性进展[8]，将测井资料与人工智能技术相结合来提升密度测井的精度已成为一种重要方法。

支持向量回归(Support Vector Machine，SVM)[9]、长短记忆网络(Long Short-Term Memory，LSTM)[10]、随机森林回归(Random Forest，RFR)[11]和梯度提升决策树(Gradient Boosting Decision Tree，GBDT)[12]等是目前拟合性能好，泛化能力强的机器学习模型。在过去二十多年实践中，梯度提升决策树(GBDT)由于其高预测性，高适应性在众多机器学习模型中脱颖而出。Chen 等人针对梯度提升决策树(GBDT)训练速度慢、时间复杂度高以及难以并行化等问题进行了改进，提出了 XGBoost，这是一个高效、灵活且具有强大可移植性的最优分布式决策梯度提升库[13]。相较之下，XGBoost 在性能和速度上都表现出色。其决策树结构直接支持模型可视化，同时通过全局特征重要性评估，提升了模型的解释性[14]。以往的研究表明，引入分解算法的混合预测模型比单一预测模型可以获得更准确的预测结果，与单一分解技术相比，使用两阶段分解技术将更充分地减弱序列的非平滑性[15]。XGBoost 作为人工智能领域比较前沿的一种算法，其收敛速度的快慢和拟合能力的强弱与多种超参数有关，如果对 XGBoost 的超参数集进行优化选择，可以充分发挥其预测性能。Ye 等将交叉验证法(LOO)与 XGBoost 算法相结合，建立了 LOO-XGBoost 预测模型来解决非线性回归问题[16]。Liu 等将粒子群优化算法(PSO)与 XGBoost 相结合，建立了 PSO-XGBoost 预测模型进行 NMR 测井响应预测[17]。彭英等将 Focal-Loss 函数与 XGBoost 结合，提出了 FL-XGBoost 模型进行岩性识别[18]。麻雀搜索算法作为一种新型群体智能优化算法，具备强大的求解能力和快速收敛的特性，已广泛应用到各类回归预测算法领域。

本文利用 SSA 辅助选择 XGBoost 的最优超参数集，然后利用 XGBoost 的非线性映射能力，建立了 SSA-XGBoost 回归预测模型对密度测井数据进行处理，提高密度测井数据的精度。并将预测结果与 XGBoost、SSA-SVR、SSA-LSTM 和 SSA-RFR 模型预测的结果进行比较，研究 SSA-XGBoost 模型在提高密度测井数据精度方面的应用。

1 SSA-XGBoost 算法原理

XGBoost 是一种可扩展的机器学习系统。它在内部实现了 GBDT 模型，是对传统的 GBDT 算法的优化[19]。该模型的复杂性得到了显著减小，可以有效解决传统模型经常出现的过拟合问题，并在保持极快运算速度的同时实现高精度。XGBoost 的预测原理是连续添加树模型，根据生长树的特征对其进行划分，并拟合上次预测的残差。最终预测值是每棵树获得的预测值之和，表达式为：

$$\hat{y}_i = \Phi\left(x_i\right) = \sum_{k=1}^{K} f_k\left(x_i\right), f_k \in F \tag{1}$$

在公式(1)中，f_k 为独立树，F 为回归树空间。

作为提高密度测井数据精度的回归预测模型，选择目标函数作为损失函数和惩罚函数的叠加：

$$L(\Phi) = \sum_i l\left(\hat{y}_i, y_i\right) + \sum_k \Omega\left(f_k\right) \tag{2}$$

其中 l 为损失函数，Ω 为惩罚函数。惩罚函数表示为：

$$\Omega(f) = \gamma T + \frac{1}{2}\lambda \| w \|^2 \tag{3}$$

在不影响精度的情况下，通过二阶泰勒展开的最终损失函数为：

$$\sum_{i}\left[l(y_i,\hat{y}_i^{(k-1)})+g_i f_k(x_i)+\frac{1}{2}h_i f_k^2(x_i)\right] \tag{4}$$

对于决策树，每次拆分都会计算一次目标函数。当从分裂中获得的信息增益成本超过极限时，决策树停止分裂，决策树的预测得分为最终的目标函数。

麻雀搜索算法(Sparrow Search Algorithm， SSA）通过模拟麻雀的进食和捕食者回避行为来优化模型参数[20]。它具有强大的优化能力和快速的收敛速度。SSA 可以有效地减少搜索空间，提高搜索效率，避免搜索空间中的局部最优解，比粒子群优化算法（PSO）和灰狼优化算法（GWO）等其他优化算法具有更好的优化能力[21-23]。SSA 算法公式如下：

$$X_{i,d}^{t+1}=\begin{cases} X_{i,d}^{t}\cdot\exp\left(\dfrac{-i}{\alpha\cdot G_{\max}}\right), R_2<ST \\ X_{i,d}^{t}+Q\cdot L, R_2\geq ST \end{cases} \tag{5}$$

在公式（5）中，X 为麻雀的位置，在算法迭代过程中表示当前解的状态；t 为迭代次数；i 为种群个体编号；d 为优化问题的变量个数；G_{max} 为最大迭代次数；α 为（0,1）分布的随机数；Q 为标准正态分布的随机数；L 为单位矩阵；ST 为警告阈值，其取值范围在[0.5,1]之间；R_2 为警告值。

设置麻雀数量为 25，最大迭代次数为 120，以配置 SSA 参数。利用 SSA 对 XGBoost 的最大迭代次数、树的深度以及学习率进行优化，以提高模型的预测性能和运算效率。之后，利用参数优化后的 SSA-XGBoost 模型对密度测井数据进行预测。

2 模型建立及数据处理

2.1 MCNP5 模拟

蒙特卡罗(Monte Carlo N-particle code, MCNP)数值模拟技术已广泛应用于核辐射测井领域，包括核测井机理和测井响应研究、仪器刻度和校正以及新方法和算法应用[24-26]。为了构建模型的数据集合和验证 SSA-XGBoost 密度预测模型各方面的性能，利用 MCNP5 建立了图 1 所示模型，仪器模型规格参数如表 1 所示。

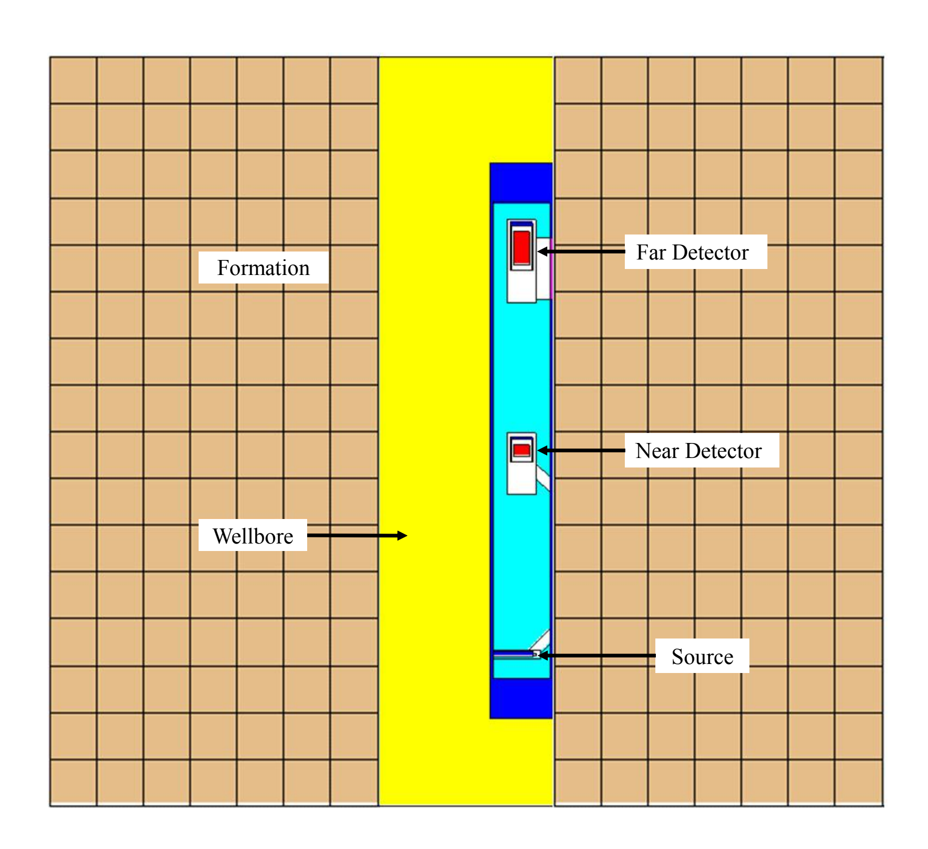


图 1　双探测器密度测井仪器模型

Fig.1　Model of the dual detector density logging tool

表 1　仪器模型参数

Table 1　Parameters of the instrument model

参数 Parameters	规格 Instructions
伽马源 Gamma ray source	Cs^{137}, 能量 Energy:0.662 Mev
探测器类型 Detector types	近探头 Near detector: 远探头 NaI, Far detector: NaI
源距 Sonde-to-source spacing	近源距 Near space: 19cm, 远源距 Far space: 39cm
仪器设计 Instrument design	直径 Diameter: 6cm, 长度 Length: 55cm

综合考虑实际情况和模拟数据的误差，设置地层填充材料为饱含水石灰岩，井眼用淡水填充，井眼直径为 15cm，该双探测器密度测井仪器放置在钻铤中，并贴井壁，模型之外的空间设置为真空。模拟的伽马光子粒子数设置为 1×10^8,探测器记录能量大于 0.15 MeV 的伽马光子通量，并且保证计数率相对误差均小于 2%。表 2 展示了模拟过程中不同密度的饱含水石灰岩长、短源距探测器计数率的统计误差。

表 2　不同探头计数率统计误差
Table 2　Statistical error of count rates for different detector

密度 Density / g•cm^{-3}	短探头计数率 Near detector count rates	短探头统计误差 Near detector statistical error	长探头计数率 Far detector count rates	长探头统计误差 Far detector statistical error
2.001	3.063E-08	0.62%	3.839E-10	1.16%
2.304	2.801E-08	0.68%	2.155E-10	1.21%
2.651	2.497E-08	0.75%	1.120E-10	1.27%
2.874	2.352E-08	0.80%	7.874E-11	1.33%
3.011	2.256E-08	0.82%	5.274E-11	1.47%

利用 MCNP5 模拟不同孔隙度和不同密度的石灰岩地层，双探测器密度测井仪器被放置在模拟地层中进行数据采样，获取对应模拟地层的长、短源距探测器计数率和地层的真实密度，以此方法做正演模拟，共获得 1000 组地层密度模拟数据。处理模拟数据量统计如表 3 所示，采用离散化处理方法，将该模型建立的密度数据以地层孔隙度变化划分为 10 组并编号。将 K1、K2、K3、K4、K6、K7、K8 和 K10 段作为机器学习模型的训练集，选择 K5 和 K9 作为测试集以评估模型的预测性能和泛化能力。

表 3　处理数据统计
Table 3　Statistical analysis of processed data

编号 Number	孔隙度 Porosity / P.U.	孔隙流体 Pore fluid	密度范围 Density range / g•cm^{-3}
K1	0	H_2O	2.46~2.71
K2	5	H_2O	2.31~2.68
K3	10	H_2O	2.51~2.94
K4	15	H_2O	2.44~2.83
K5	20	H_2O	2.36~2.88
K6	25	H_2O	2.15~2.87
K7	30	H_2O	2.21~2.98
K8	35	H_2O	2.14~2.91
K9	40	H_2O	2.25~2.95
K10	45	H_2O	2.41~2.96

2.2 XGBoost 模型优化

为了检验所构建的用于提高密度测井数据精度的 SSA-XGBoost 模型是否有效，先将 XGBoost 用于该密度精度回归预测模型。在考虑到模型精度，减少过拟合以及算法效率等参数影响后，利用 SSA 算法优化最大迭代次数，树的深度和学习率，并将该模型与 XGBoost 模型进行比较。最终优化结果如表 4 所示：

表 4　参数优化结果
Table 4　Results of parameter optimization

模型 Model	最大迭代次数 Maximum iterations	学习率 Learning rate	树的深度 Tree depth
SSA-XGBoost	90	0.9182	6
XGBoost	81	0.9546	11

根据上述步骤，利用 SSA-XGBoost 模型对测试集中的密度测井数据进行预测，并将预测结果与 XGBoost 模型和模拟地层真实密度作对比,SSA-XGBoost 模型预测结果的准确度得到了一定程度的提升，数据结果如图 2 所示：

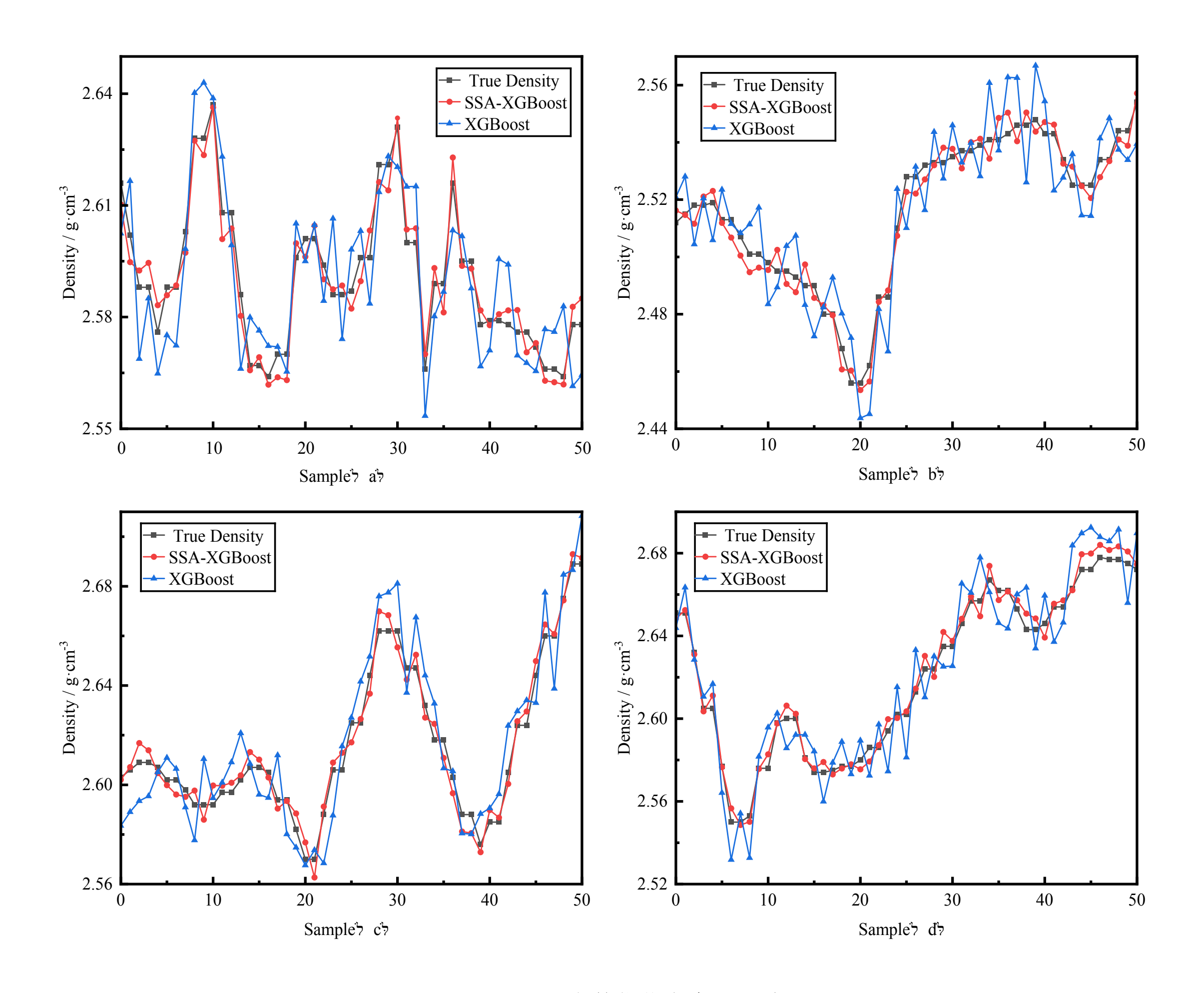


图 2　SSA-XGBoost 参数优化密度预测结果对比

Fig.2　Comparison of SSA-XGBoost parameter optimization density prediction results

为了进一步比较 SSX-XGBoost 模型与 XGBoost 模型的预测性能，选取了测试集中的 200 个数据点进行模型误差分析，SSA-XGBoost 模型的预测精度得到了有效的提高。图 3 展示了 SSA-XGBoost 模型和 XGBoost 模型在不同测试集的回归预测密度测井曲线，可以明显的看出经过多个超参数优化的 SSA-XGBoost 模型总体预测性能较好，预测更准确，能更好的反应地层的真实密度。这是因为 SSA-XGBoost 能够分离数据的不同层析结构，消除潜在的冗余信息，降低过拟合并提升模型的泛化能力，同时该算法能够生成更高质量的特征来反映时间序列的内在动态，从而在后续的训练模型中得到更为准确的预测结果。

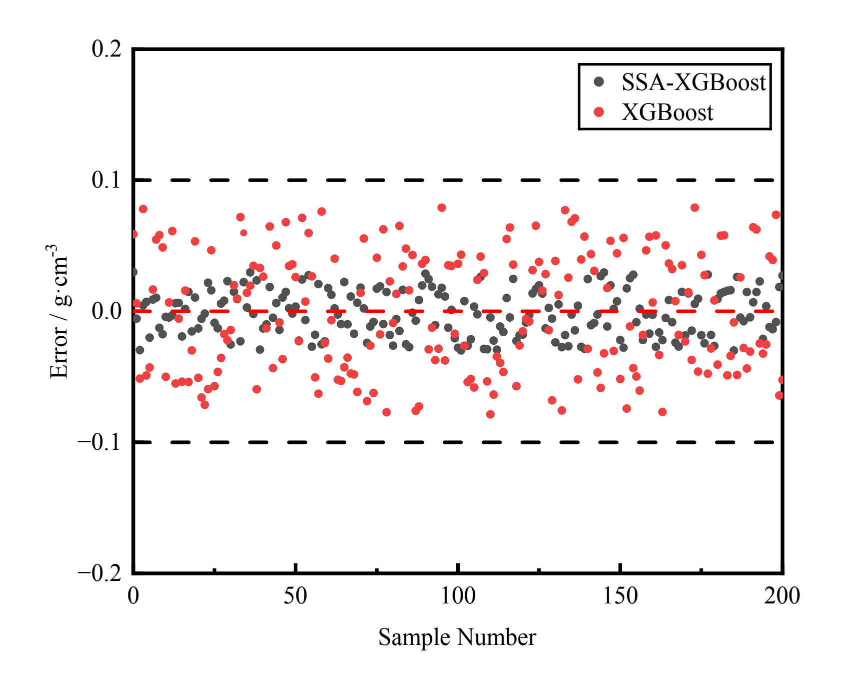


图 3　SSA-XGBOOST 超参数优化模型误差对比

Fig.3　Comparison of SSA-XGBOOST model errors for hyperparameter optimization

2.3 模拟数据预测结果对比与分析

到目前为止，在传统的密度测井补偿方法中，脊-肋图法仍是使用最广泛的方法[27]。在使用脊-肋图法进行密度校正时，将“脊线”看作是直线，“肋线”也作近似线性处理[28]，则补偿密度测井的基本公式为：

$$\rho_b = \frac{1}{A_L}\left(\ln N_L - B_L\right) + a\left[\frac{1}{A_L}(\ln N_L - B_L) - \frac{1}{A_S}(\ln N_S - B_S)\right] \tag{6}$$

在公式（6）中，A_L、A_S分别为长、短距探头对地层的灵敏度，B_L、B_S为与光子通量有关的常数，N_L、N_S分别为长短源距探头所对应的计数率，a 为与仪器结构系数有关的常数。

双对数坐标下的脊-肋图版可以对井眼条件下不同厚度的泥饼进行校正，描述密度测井曲线的性能，简化密度测井数据的处理过程[29]。在井眼环境比较复杂的情况下，比如受重晶石，赤铁矿等影响严重，“肋线”可能是二次或者多次曲线，为了得到更精确的处理结果，可以采用二次或者多次拟合的方法，尽管这样可以在一定程度上对密度进行补偿，但在实际情况下，计算出的密度值仍然具有较大的偏差，不能用来反映地层的真实密度。

为评价“脊-肋”图法和 SSA-XGBoost 模型的密度补偿性能，同时采用 SSA 算法对支持向量回归(SVR)、随机森林(RFR)、长短期记忆网络(LSTM)超参数优化，得到预测性能更好的 SSA-SVR、SSA-RFR 和 SSA-LSTM 模型。SSA-XGBoost 和 XGBoost 属于优化的梯度提升决策树算法，SSA-SVR 属于监督机器学习算法，SSA-RFR 属于集成学习算法，SSA-LSTM 属于循环神经网络算法。将 MCNP5 模拟所得密度测井数据和长、短源距探测器计数率作为训练模型的输入参数，分别建立回归预测模型，以探究不同模型对密度精度的影响。图 4 展示了不同模型的密度精度回归预测图，横轴为地层的真实密度，纵轴为不同模型预测的地层密度，数据点越接近对角线，表明该模型预测地层密度性能越好。

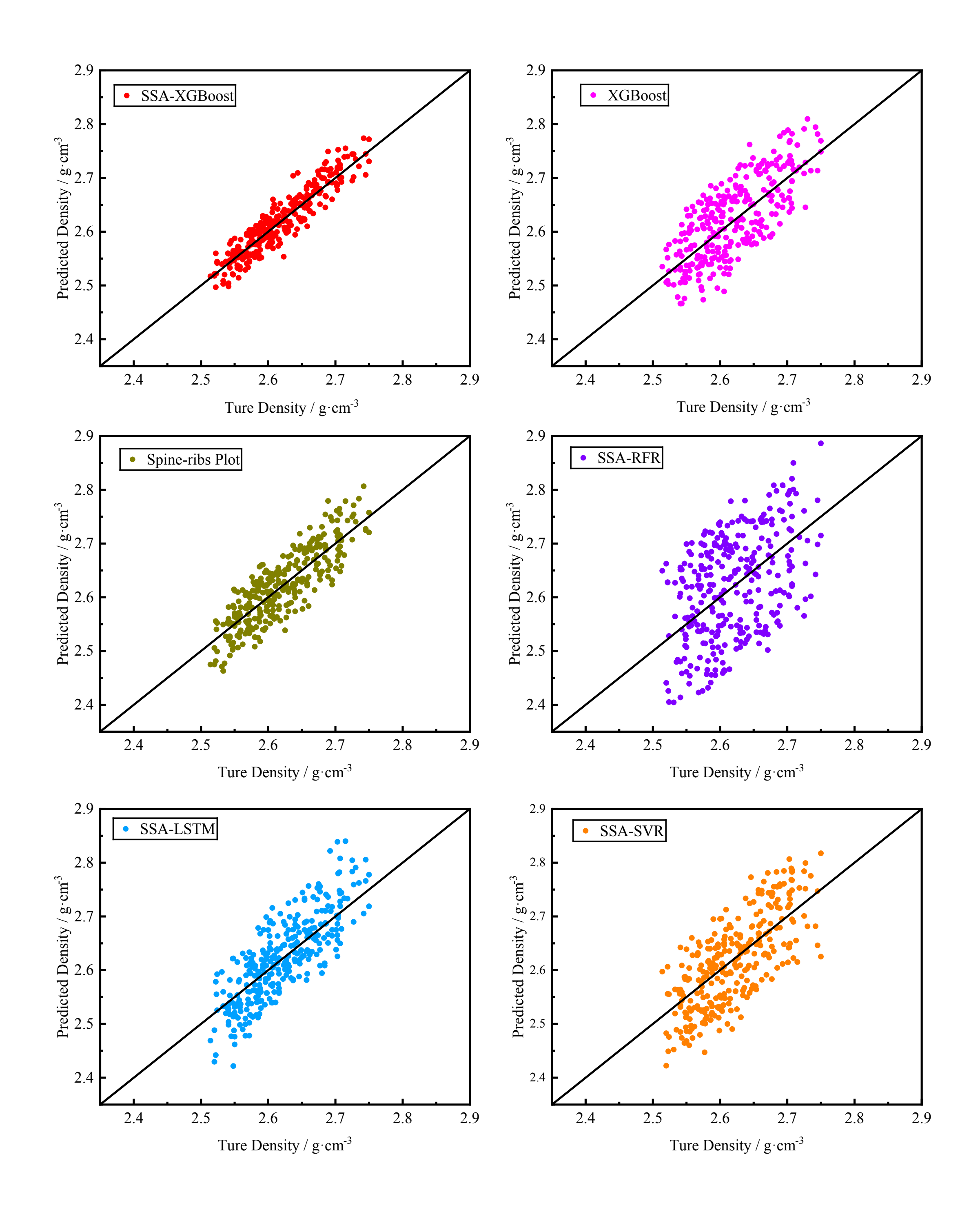


图 4　不同模型模拟数据密度预测结果对比

Fig.4　Comparison of simulation data density prediction results of different models

为了进一步比较各个模型性能的差异，引入评价指标量化评估模型性能，用评价指标来衡量预测密度与地层真实密度之间的差异。本次研究选用了 4 个评价指标：平均绝对误差(MAE)、均方根误差(RMSE)、标准差(SD)以及预测模型拟合准确度的决定系数(R^2)。计算公式如下所示：

$$\mathrm{MAE}=\frac{1}{m}\sum_{i=1}^{m}\left|\hat{y}_i-y_i\right| \tag{7}$$

$$\mathrm{RMSE}=\sqrt{\frac{1}{m}\sum_{i=1}^{m}(\hat{y}_i-y_i)^2} \quad (8)$$

$$\mathrm{SD}=\sqrt{\frac{\sum_{i=1}^{m}(\mathrm{y}_i-\overline{y}_\mathrm{i})^2}{\mathrm{m}}} \quad (9)$$

$$R^2=1-\frac{\sum_{i=1}^{m}(\hat{y}_i-y_i)^2}{\sum_{i=1}^{m}(\overline{y}-y_i)^2} \quad (10)$$

在上述公式中，m为测试集数据点数量，$\hat{y}_i$为模型预测值，y_i为对应地层的真实密度，$\overline{y}_\mathrm{i}$为预测地层密度均值，$\overline{y}$为地层的真实密度均值。

利用各个回归预测模型和脊-肋图模型完成对训练井数据的训练后，将其应用到 K5 和 K9 测试集中，共包含 200 个数据组，表 5 列出了 6 种模型对测试井密度进行预测的评价指标数据。SSA-XGBoost 回归预测模型表现出良好的预测效果，精度较高，SSA-RFR 模型预测效果相对较差，此结论与上述不同模型回归实验所得结论相对应。从平均绝对误差的角度分析发现，SSA-XGBoost 模型预测地层密度的平均绝对误差约为 0.0174 g•cm^{-3}，脊-肋图模型预测地层密度时的平均绝对误差达到了 0.0284 g•cm^{-3}，两者的密度精度具有较大差别，说明对于 XGBoost 模型来说，尽管它的性能已经优于其它回归预测模型，但通过 SSA 的超参数优化可以更显著的提高模型的预测性能。脊-肋图版法在这两口测试井所得密度的平均绝对误差和均方根误差分别为 0.0284 g•cm^{-3} 和 0.0346 g•cm^{-3}，说明传统的提高密度精度的方法在遇到井眼环境复杂和岩性多样的地层时，会产生不同程度的误差，需要进行井眼校正和各种参数校正。

表 5　不同模型评价指标
Table 5　Evaluation metrics for different models

Model	SSA-XGBoost	Spine-ribs Plot	XGBoost	SSA-LSTM	SSA-RFR	SSA-SVR
MAE	0.0174	0.0284	0.0314	0.0330	0.0721	0.0459
RMSE	0.0206	0.0346	0.0365	0.0382	0.0834	0.0530
STD	0.0613	0.0682	0.0744	0.0789	0.0975	0.0835
R^2	0.9485	0.8182	0.6448	0.7499	-0.1914	0.5187

通过对表 5 中各模型误测结果的误差分析，在评估各个模型的准确性和可靠性时，为了直观且有效的对比不同模型预测结果与真实数据之间的相似度与差异，采用泰勒图将多种评价指标简化到二维图上，以紧凑且易于理解的方式展现各个模型的综合性能[30]。泰勒图能根据模型预测结果的相关性系数、均方根误差和标准差来评价模型的性能[31]。基本公式如下：

$$R\mathrm{MSE}=S\mathrm{D}_\mathrm{t}^2+S\mathrm{D}_\mathrm{p}^2-2S\mathrm{D}_\mathrm{t}S\mathrm{D}_\mathrm{p}\mathrm{R}^2 \quad (11)$$

式中，$S\mathrm{D}_\mathrm{t}$ 和 SD_p 分别为地层真实密度和模型预测值的标准差。

令 $\mathrm{R}^2=\cos\theta$，则得到与 $R\mathrm{MSE}$、$S\mathrm{D}_\mathrm{t}$ 和 SD_p 相关的余弦公式：

$$R\mathrm{MSE}=S\mathrm{D}_\mathrm{t}^2+S\mathrm{D}_\mathrm{p}^2-2S\mathrm{D}_\mathrm{t}S\mathrm{D}_\mathrm{p}\cos\theta \quad (12)$$

利用公式（12）将不同模型的误差参数转化为极坐标，将其投影到坐标系中，如图 5 所示。其中红色坐标轴表示标准差，黄色弧线表示均方根误差，越靠近圆弧中心店，误差越小，蓝色刻度线表示相关系数，越接近 1 则相关性越好。综合考虑这些预测模型和脊-肋图误差计算结果的比较，SSA-SVR 模型预测地层密度效果较差，SSA-LSTM 预测地层密度精度一般，与地层真实密度相比存在一定较大误差。脊-肋图模型可以较为准确的预测地层密

度，与地层真实密度的误差较小。SSA-XGBoost 模型的预测性能优于脊-肋图模型，预测结果的准确度进一步得到提高，该方法在提高密度测井数据精度方面具有很大潜力。

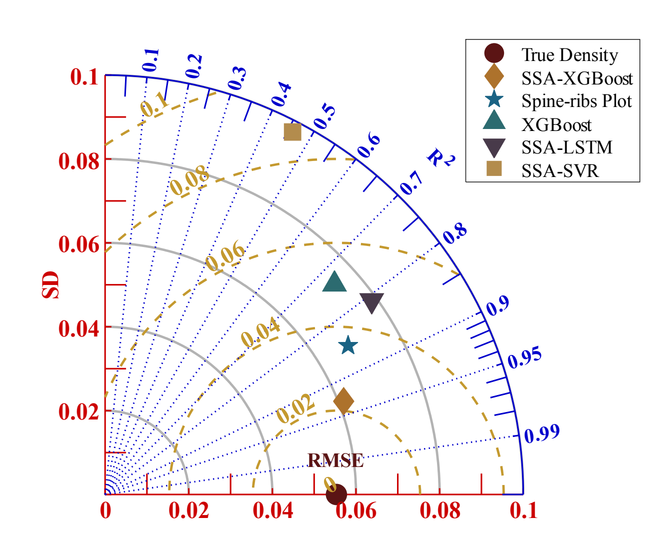


图 5　模型性能对比泰勒图

Fig.5　Model performance comparison Taylor plot

2.4 实测数据处理对比及分析

为验证 SSA-XGBoost 模型的准确性和泛化能力，在 MCNP5 数值模拟 1000 组密度测井数据的基础上，将在密度预测对比中泛化能力表现最好的 SSA-XGBoost 模型和传统的脊-肋图模型进行对实例测井数据进行处理，选取了实际应用井中不同层段的测井数据进行实验结果的可视化，结果如图 6 所示。其展示了 SSA-XGBoost 模型与脊-肋图法对珠江口盆地白云构造区内的某口测井深度分别为 4415~4460m 和 4970~5015m 的油气井可视化预测结果。在图中，深度值左侧曲线为常规测井资料解释曲线，深度值右侧的曲线展示了 SSA-XGBoost 模型与脊-肋图法预测密度曲线和真实密度曲线 DEN 的对比情况以及分别对应的密度误差曲线。从图中可以看出本文提出的 SSA-XGBoost 模型预测值与真实值误差较小，呈现出较高的密度精度，同时也验证了该方法的有效性和精确性。

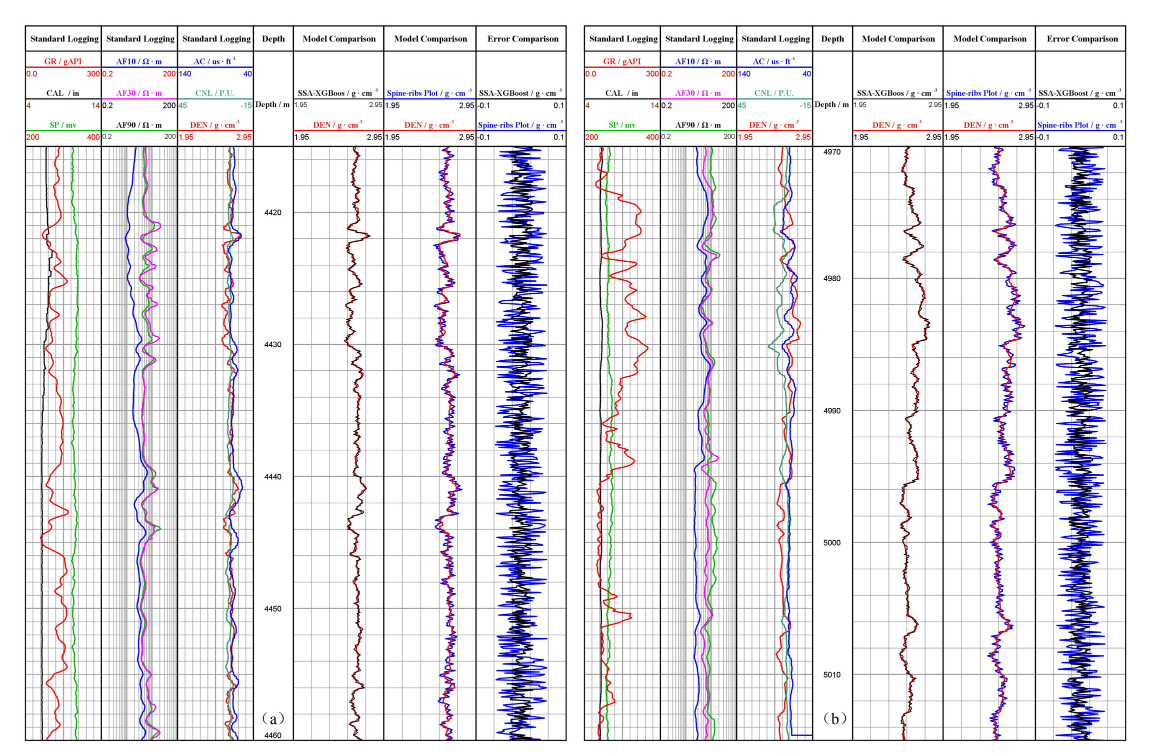


图 6　4415~4460m 实测数据处理对比 4415~4460m(a)和 4970~5015m(b)

Fig.6. Comparison of measured data processing of 4415~4460m 4415~4460m(a) and 4970~5015m(b)

3 结语

1）本文提出了一种基于 SSA-XGBoost 的密度测井数据高精度预测方法。利用蒙特卡罗数值模拟所得 1000 组密度测井数据作为该模型的训练集和测试集，将多种机器学习方法与传统脊-肋图法所得结果与之对比，分析其应用效果。

2）模拟数据处理结果表明，利用 SSA 算法对 XGBoost 模型进行超参数优化，能显著提高其对地层密度的预测性能。SSA-XGBoost 模型表现出更强的泛化能力，能够有效解决回归预测中的过拟合问题，提高密度预测精度。

3）与其他机器学习模型（如 XGBoost、SSA-LSTM、SSA-SVR、SSA-RFR）和传统方法（如脊-肋图法）相比，SSA-XGBoost 模型在对模拟数据进行回归预测时表现出更高的预测精度。其预测性能在各项评价指标上均优于其他方法，预测测试集地层密度平均绝对误差仅有 $0.0174g \cdot cm^{-3}$，远低于脊-肋图法的误差 $0.0284g \cdot cm^{-3}$。利用 SSA-XGBoost 模型处理实际密度测井数据时仍具有较高的预测精度，该方法可以有效解决其他模型在预测地层密度时精度较低的问题。

作者贡献声明　李瑞负责数值模拟、数据处理、机器学习算法优化和应用及论文编写与修改工作；吴文圣负责文章思路的提出和指导。

参考文献

1 Tittman J, Wahl J S. The physical foundations of formation density logging (gamma-gamma)[J]. Geophysics, 1965, 30(2): 284-294. DOI: 10.1190/1.1439574.

2 Zhang Q, Wang Y. Formation density and photoelectric index calculation using an ultra-slim density tool[J]. Geoenergy Science and Engineering, 2024: 212817. DOI: 10.1016/j.geoen.2024.212817.

3 赵太平, 刘天定, 王秀明. 补偿密度测井仪器刻度对测井响应的影响分析[J]. 测井技术, 2011, 35(6): 576-580. DOI: 10.16489/j.issn.1004-1338.2011.06.019.

Zhao, T., Liu, T., & Wang, X. (2011). Analysis of the impact of calibration on logging responses for compensated density logging instruments. Well Logging Technology, 35(6), 576-580. DOI: 10.16489/j.issn.1004-1338.2011.06.019.

4 吴文圣, 黄隆基. 三探测器密度测井的 Monte Carlo 模拟[J]. 地球物理学报, 2004, 47(1): 164-170.

Wu, W., & Huang, L. (2004). Monte Carlo simulation of three-detector density logging. Chinese Journal of Geophysics, 47(1), 164-170.

5 李崇儒, 陈国伟. 补偿密度测井仪的刻度与密度计算[J]. 煤田地质与勘探, 1992, 20(4): 52-58.

Li, C., & Chen, G. (1992). Calibration and density calculation of compensated density logging instrument. Coal Geology & Exploration, 20(4), 52-58.

6 Limin Z, Zhenyan C, Yongcheng C. Application of Multi-Parameter Lithologic Seismic Inversion to Predict Glutenite Reservoirs in Santaizi Subsag of Damintun Sag[J]. China Petroleum Exploration, 2007, 12(4): 53. DOI: 10.6038/pg20130447.

7 叶义平, 盛云霞, 冯丽萍, 等. 横向测井资料的孔隙度测井解释方法研究[J]. 石油天然气学报, 2005, 27(3): 348-350.

Ye, Y., Sheng, Y., Feng, L., et al. (2005). Study on porosity logging interpretation method of lateral logging data. Journal of Oil and Gas Technology, 27(3), 348-350.

8 程希, 周军, 傅海成, 等. 机器学习算法在地球物理测井中的适用性及应用[J]. 西北地质, 2023, 56(4): 336-348. DOI: 10.12401/j.nwg.2023029.

Cheng, X., Zhou, J., Fu, H., et al. (2023). Applicability and application of machine learning algorithms in geophysical logging. Northwestern Geology, 56(4), 336-348. DOI: 10.12401/j.nwg.2023029.

9 Zhang Z, Zhang H, Li J, et al. Permeability and porosity prediction using logging data in a heterogeneous dolomite reservoir: An integrated approach[J]. Journal of Natural Gas Science and Engineering, 2021, 86: 103743. DOI: 10.1016/j.jngse.2020.103743.

10 Yu Y, Si X, Hu C, et al. A review of recurrent neural networks: LSTM cells and network architectures[J]. Neural computation, 2019, 31(7): 1235-1270. DOI: 10.1162/neco_a_01199.

11 郭建宏, 张占松, 张超谟, 等. 用地球物理测井资料预测煤层气含量——基于斜率关联度—随机森林方法的工作案例[J]. 物探与化探, 2021 (1): 18-28. DOI: 10.11720/wtyht.2018.0177.

Guo, J., Zhang, Z., Zhang, C., et al. (2021). Predicting coalbed methane content using geophysical logging data: A case study based on the slope correlation and random forest methods. Geophysical and Geochemical Exploration, (1), 18-28. DOI: 10.11720/wtyht.2018.0177.

12 Subasi A, El-Amin M F, Darwich T, et al. Permeability prediction of petroleum reservoirs using stochastic gradient boosting regression[J]. Journal of Ambient Intelligence and Humanized Computing, 2022: 1-10. DOI: 10.1007/s12652-020-01986-0.

13 Shwartz-Ziv R, Armon A. Tabular data: Deep learning is not all you need[J]. Information Fusion, 2022, 81: 84-90. DOI: 10.1016/j.inffus.2021.11.011.

14 Pan S, Zheng Z, Guo Z, et al. An optimized XGBoost method for predicting reservoir porosity using petrophysical logs[J]. Journal of Petroleum Science and Engineering, 2022, 208: 109520. DOI: 10.1016/j.petrol.2021.109520.

15 Gao Z, Yin X, Zhao F, et al. A two-layer SSA-XGBoost-MLR continuous multi-day peak load forecasting method based on hybrid aggregated two-phase decomposition[J]. Energy Reports, 2022, 8: 12426-12441. DOI: 10.1016/j.egyr.2022.09.008.

16 Xu B, Tan Y, Sun W, et al. Study on the prediction of the uniaxial compressive strength of rock based on the SSA-xgboost model[J]. Sustainability, 2023, 15(6): 5201. DOI: 10.3390/su15065201.

17 谷宇峰, 张道勇, 鲍志东. 测井资料 PSO-XGBoost 渗透率预测[J]. 石油地球物理勘探, 2021, 56(1): 26-37. DOI: 10.13810/j.cnki.issn.1000-7210.2021.01.003.

Gu, Y., Zhang, D., & Bao, Z. (2021). Permeability prediction using logging data and PSO-XGBoost. Oil Geophysical Prospecting, 56(1), 26-37. DOI: 10.13810/j.cnki.issn.1000-7210.2021.01.003.

18 彭英, 李克文, 朱应科, 等. 基于 FL-XGBoost 算法的砂泥岩识别方法——以胜利油田牛庄地区为例[J]. 油气地质与采收率, 2023, 30(1): 76-85. DOI: 10.13673/j.cnki.cn37-1359/te.202205031.

Peng, Y., Li, K., Zhu, Y., et al. (2023). Identification method of sandstone and mudstone based on FL-XGBoost algorithm: A case study of Niuzhuang area in Shengli Oilfield. Petroleum Geology and Recovery Efficiency, 30(1), 76-85. DOI: 10.13673/j.cnki.cn37-1359/te.202205031.

19 谈林涛, 李军良, 任昺, 等. 基于 RB-XGBoost 算法的智能电网调度控制系统健康度评价模型[J]. Electric Power Automation Equipment/Dianli Zidonghua Shebei, 2020, 40(2). DOI: 10.16081/j.epae.202001024.

Tan, L., Li, J., Ren, B., et al. (2020). Health evaluation model of smart grid dispatch control system based on RB-XGBoost algorithm. Electric Power Automation Equipment, 40(2). DOI: 10.16081/j.epae.202001024.

20 Xue J, Shen B. A novel swarm intelligence optimization approach: sparrow search algorithm[J]. Systems science & control engineering, 2020, 8(1): 22-34. DOI: 10.1080/21642583.2019.1708830.

21 Liu Y, Shi Y, Chen H, et al. Chaos-assisted multi-population salp swarm algorithms: Framework and case studies[J]. Expert Systems with Applications, 2021, 168: 114369. DOI: 10.1016/j.eswa.2020.114369.

22 Niu P, Niu S, Chang L. The defect of the Grey Wolf optimization algorithm and its verification method[J]. Knowledge-Based Systems, 2019, 171: 37-43. DOI: 10.1016/j.knosys.2019.01.018.

23 Marini F, Walczak B. Particle swarm optimization (PSO). A tutorial[J]. Chemometrics and Intelligent Laboratory Systems, 2015, 149: 153-165. DOI: 10.1016/j.chemolab.2015.08.020.

24 Brown F B, Barrett R F, Booth T E, et al. MCNP version 5[J]. Trans. Am. Nucl. Soc, 2002, 87(273): 02-3935.

25 Solberg T D, DeMarco J J, Chetty I J, et al. A review of radiation dosimetry applications using the MCNP Monte Carlo code[J]. Radiochimica Acta, 2001, 89(4-5): 337-355. DOI: 10.1524/ract.2001.89.4-5.337.

26 Gardner R P, Xu L. Status of the Monte Carlo library least-squares (MCLLS) approach for non-linear radiation analyzer problems[J]. Radiation Physics and Chemistry, 2009, 78(10): 843-851.

27 孙培伟, 张建民, 岳爱忠, 等. 双源距密度测井的蒙特卡罗数值模拟[J]. 原子能科学技术, 2007, 41(5): 595-600.

Sun, P., Zhang, J., Yue, A., et al. (2007). Monte Carlo numerical simulation of dual-source compensated density logging. Atomic Energy Science and Technology, 41(5), 595-600.

28 Huawei Y, Jianmeng S, Xiaopeng L, et al. Monte Carlo simulation on influence of cutting beds on density log in the horizontal wells[J]. Nuclear Techniques, 2009, 32.

29 黄隆基. 放射性测井原理[M]. 石油工业出版社, 1985:81~99.

Huang, L. (1985). Principles of Radioactive Logging. Petroleum Industry Press, 81-99.

30 Pincus R, Batstone C P, Hofmann R J P, et al. Evaluating the present‐day simulation of clouds, precipitation, and radiation in climate models[J]. Journal of Geophysical Research: Atmospheres, 2008, 113(D14). DOI: 10.1029/2007JD009334.

31 Xu B, Tan Y, Sun W, et al. Study on the prediction of the uniaxial compressive strength of rock based on the SSA-gboost model[J]. Sustainability, 2023, 15(6): 5201. DOI: 10.3390/su15065201.